Explainable artificial intelligence enhances the ecological interpretability of black-box species distribution models

Ryo, Masahiro

2021-02

Ryo , M , Angelov , B , Mammola , S , Kass , J M , Benito , B M & Hartig , F 2021 , ' Explainable artificial intelligence enhances the ecological interpretability of black-box species distribution models ' , Ecography , vol. 44 , no. 2 , pp. 199-205 . https://doi.org/10.1111/ecog.05360

http://hdl.handle.net/10138/326142
https://doi.org/10.1111/ecog.05360

ECOGRAPHY

Forum

Explainable artificial intelligence enhances the ecological interpretability of black-box species distribution models

Masahiro Ryo, Boyan Angelov, Stefano Mammola, Jamie M. Kass, Blas M. Benito and Florian Hartig

M. Ryo (https://orcid.org/0000-0002-5271-3446) ✉ (masahiroryo@gmail.com), Inst. of Biology, Freie Univ. Berlin, Berlin, Germany, and Berlin-Brandenburg Inst. of Advanced Biodiversity Research (BBIB), Berlin, Germany, and Leibniz Centre for Agricultural Landscape Research (ZALF), Müncheberg, Germany. – B. Angelov, Association for Computing Machinery (ACM), New York, USA. – S. Mammola (https://orcid.org/0000-0002-4471-9055), Molecular Ecology Group (MEG), Water Research Inst. (IRSA), National Research Council (CNR), Verbania Pallanza, Italy, and Laboratory for Integrative Biodiversity Research (LIBRe), Finnish Museum of Natural History (LUOMUS), Univ. of Helsinki, Helsinki, Finland. – J. M. Kass (https://orcid.org/0000-0002-9432-895X), Biodiversity and Biocomplexity Unit, Okinawa Inst. of Science and Technology Graduate Univ., Kunigami-gun, Okinawa, Japan. – B. M. Benito (https://orcid.org/0000-0001-5105-7232), Dept of Ecology and Multidisciplinary, Inst. for Environment Studies 'Ramon Margalef', Univ. of Alicante, San Vicente del Raspeig, Alicante, Spain. – F. Hartig, Theoretical Ecology, Faculty of Biology and Pre-Clinical Medicine, Univ. Regensburg, Regensburg, Germany.

Ecography
44: 199–205, 2021
doi: 10.1111/ecog.05360

Subject Editor: Thiago Rangel
Editor-in-Chief: Miguel Araújo
Accepted 27 October 2020

Species distribution models (SDMs) are widely used in ecology, biogeography and conservation biology to estimate relationships between environmental variables and species occurrence data and make predictions of how their distributions vary in space and time. During the past two decades, the field has increasingly made use of machine learning approaches for constructing and validating SDMs. Model accuracy has steadily increased as a result, but the interpretability of the fitted models, for example the relative importance of predictor variables or their causal effects on focal species, has not always kept pace. Here we draw attention to an emerging subdiscipline of artificial intelligence, explainable AI (xAI), as a toolbox for better interpreting SDMs. xAI aims at deciphering the behavior of complex statistical or machine learning models (e.g. neural networks, random forests, boosted regression trees), and can produce more transparent and understandable SDM predictions. We describe the rationale behind xAI and provide a list of tools that can be used to help ecological modelers better understand complex model behavior at different scales. As an example, we perform a reproducible SDM analysis in R on the African elephant and showcase some xAI tools such as local interpretable model-agnostic explanation (LIME) to help interpret local-scale behavior of the model. We conclude with what we see as the benefits and caveats of these techniques and advocate for their use to improve the interpretability of machine learning SDMs.

Keywords: ecological modeling, explainable artificial intelligence, habitat suitability modeling, interpretable machine learning, species distribution model, xAI

Explaining and predicting where and why species occur in space and time is central to ecology, biogeography and conservation biology (Pecl et al. 2017, Araújo et al. 2019). Species distribution models (SDMs) are currently the most widely used approach for

NORDIC SOCIETY OIKOS

www.ecography.org

this purpose. SDMs estimate relationships between species occurrence data and environmental variables such as land use and climatic factors. These models can be used to identify factors that predict species' presence or habitat suitability and to project distributional shifts in response to environmental change (Elith and Leathwick 2009, Booth et al. 2014).

Since the first SDM applications in the early 1980s (Box 1981, Booth et al. 2014, Booth 2018), the field has steadily moved from simple statistical models (e.g. logistic regressions) to more complex statistical methods, often adopting principles or algorithms from the field of machine learning (Phillips et al. 2006, Elith and Leathwick 2009). Moreover, the community has put substantial efforts into making SDMs more easy-to-use by streamlining the model-building and analytical processes through various software packages, for example, graphical user interfaces (Scachetti-Pereira 2002, Phillips et al. 2006, de Souza Muñoz et al. 2011, Kass et al. 2018) and programming frameworks (cf. > 10 R packages available for SDMs as reviewed in Angelov 2019). With these developments, SDMs have matured into a widely applied ecological modeling tool that has resulted in more than 6000 studies using or referencing them in the past two decades (Araújo et al. 2019).

Whereas the wide availability of complex machine learning algorithms has allowed users to build more accurate SDMs, it has not necessarily enhanced their understanding of the resulting models. How and why machine learning algorithms make their predictions is often difficult to understand, which is why they are frequently referred to as 'black-box' models. In general, there is a trade-off between the accuracy and interpretability of statistical models (Breiman 2001a). Achieving both simultaneously is challenging (Guisan and Thuiller 2005), but most researchers would agree that an ideal SDM should be both accurate and easy to interpret (Phillips et al. 2004, Austin 2007, Merow et al. 2014, Halvorsen et al. 2015). It is reasonable to ask whether ecologists should sacrifice interpretability by using excessively complex algorithms for constructing SDMs in order to procure slight advantages in predictive accuracy (Qiao et al. 2015, Araújo et al. 2019).

The dilemma of gaining accuracy only at the expense of interpretability is not unique to ecology. Fields as diverse as financial risk assessment, medicine or criminal justice have recently also realized that, although machine learning algorithms have desirable properties for making accurate predictions, it is difficult to understand the rationale underlying these predictions. The lack of interpretability makes these models less reliable or acceptable for scientists and stakeholders alike (Ribeiro et al. 2016, Meske and Bunde 2020). This problem has led to the emerging research area of explainable artificial intelligence (xAI), a subfield of AI also termed interpretable machine learning (Murdoch et al. 2019), that aims at developing tools for enhancing the interpretability of complex algorithms (Carvalho et al. 2019). The field of xAI has been developing quickly in recent years, and many new methods have been proposed, reviewed and applied in various scientific fields recently (Molnar 2019, Murdoch et al. 2019, Boehmke and Greenwell 2020, Lucas 2020).

The purposes of this forum article are to provide a brief introduction to the field and several techniques of xAI and to suggest for the first time its potential applicability to SDM research (Fig. 1). This work builds upon previous studies and software that improved accessibility and understanding for novel ML tools in ecology (Cutler et al. 2007, Elith et al. 2008, 2011, Olden et al. 2008, Elith and Graham 2009, Merow et al. 2013, Ryo and Rillig 2017, Kass et al. 2018). We acknowledge that some of these methods are already routinely used, and substantial efforts have already been made to improve the interpretation of fitted machine learning models in SDM research and ecology, independently of the emergence of xAI: e.g. bootstrap approach for key variable detection (Olden and Jackson 2002), novel higher-order interaction discovery (Ryo et al. 2018), and Maxent's 'Explain' tool and multiple variable importance metrics (Phillips 2017). However, these efforts are now being rapidly synthesized and expanded in the scientific domain of xAI, and several tools are readily available. Thus, we call for attention to the tools and principles developed in this field for ecological applications.

Following the classification in Murdoch et al. (2019), interpretability methods are largely categorized into model-based interpretability and post-hoc (i.e. after model fitting) interpretability (Fig. 1; note that different terminology may also be used for method classification in other studies because the xAI domain is still at emergence and dynamic). Model-based methods employ relatively simple algorithms so that the model architecture is more easily understandable for the user (e.g. decision tree), while post-hoc methods analyzes the behaviors of fitted complex models (e.g. random forests) and explains, for example, which variables are important for predictions and how predictor variables are associated with response variables. Many post-hoc methods are applicable for analyzing any kind of complex model (model-agnostic), although some of them are only for particular models (model-specific) (Adadi and Berrada 2018, Molnar 2019). Post-hoc methods are used for understanding what the model learned from the dataset (global level) or for understanding the rationales that the model gives for each prediction (local level). In Table 1, we introduce several post-hoc model-agnostic methods at the global and local levels. These methods are not mutually exclusive, and thus can be used jointly for interpreting the same model for different purposes.

As an example from these methods, we describe the mechanistic details of the local interpretable model-agnostic explanation (LIME), a post-hoc interpretation method proposed by Ribeiro et al. (2016), because it has gained substantial attention in the AI scientific community and contributed to increase the popularity of xAI. The aim of LIME is to explain how the fitted complex model creates a prediction for a given instance (i.e. a grid cell or other local neighborhood). To this end, for each instance, LIME fits a 'local surrogate' model (a simple model; e.g. a logistic regression or decision tree) that approximates the behavior of the complex model for a limited area of the n-dimensional space defined by the predictor variables. Searching for the local surrogate model is formulated

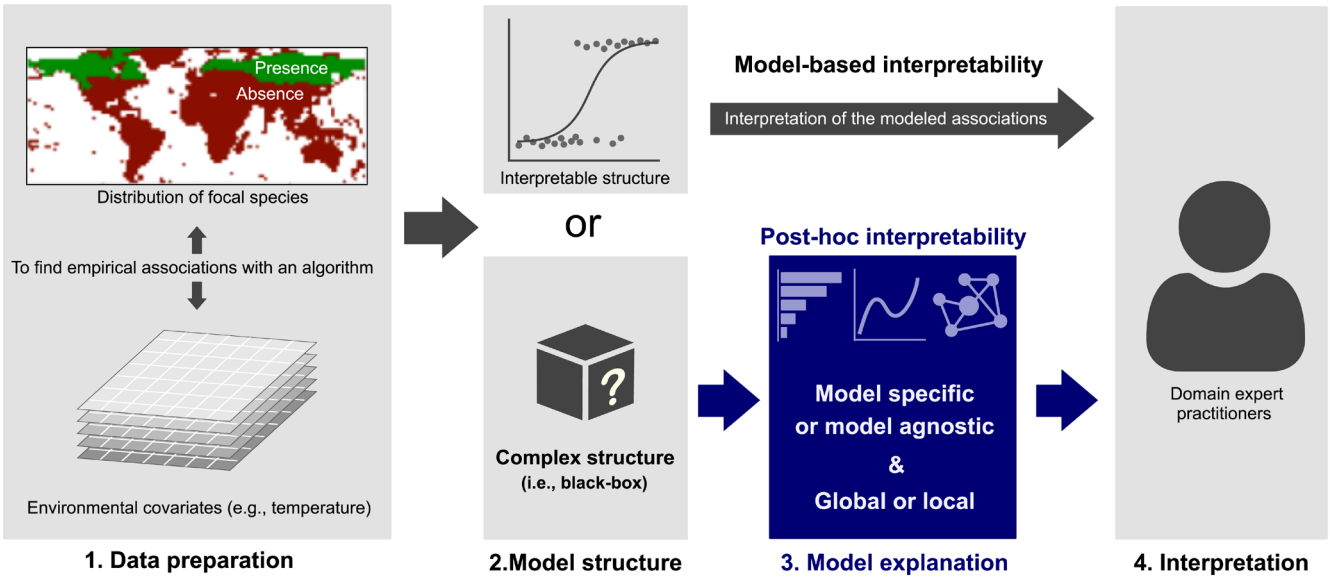

Figure 1. The role of explainable artificial intelligence (xAI) in species distribution modeling. Interpretable machine learning methods either target a direct understanding of model architecture (i.e. model-based interpretability) or interpret the model by analyzing the model behavior (post-hoc interpretability). Post-hoc interpretability methods are either model-specific (used for a specific model) or model-agnostic, meaning that they can be used for any model. Moreover, post-hoc interpretability methods are classified according to scale, global or local (global: for interpreting what the model learned from the entire variable space; local: for interpreting how each prediction is made based on the values at the instance).

Table 1. Model-agnostic post-hoc methods in explainable artificial intelligence (xAI), their approaches and potential use for species distribution models (SDMs). Model-agnostic means that they can be used for any model. Note that the list may not fully cover all available methods. For the 'level' column, 'local' means that the method is applicable for understanding how each prediction is made, while 'global' means that it is used for understanding the model learned from the dataset.

Level	Method	Approach	Utility for SDM modellers	Ref.
Local	Local interpretable model-agnostic explanation	Explains how the model predicts at a given instance with a simple parametric or machine learning model	Inspects which features are important for the prediction at a given location.	(Ribeiro et al. 2016)
Local	Shapley values	Explains the relative contribution of each feature to the prediction at a given instance, based on cooperative game theory	Inspects which features are important for the prediction at a given location.	(Lundberg and Lee 2017)
Local	Anchors	Builds if-then rules with some features and their thresholds such that the model does not change the prediction at a given instance	Identifies which features are important and how sensitive the prediction is against changes in the selected features at a given location.	(Ribeiro et al. 2018)
Local	iBreakDown	Identifies interactions and measures their contributions	Inspects which feature interactions are important for the prediction at a given location.	(Gosiewska and Biecek 2020)
Global	Permutation importance; multi-level block permutation	Measures the relative contribution of each feature to the model accuracy. The latter accounts for the covariance structure.	Quantifies which variables are important, including the mean and confidence intervals, for the accuracy of the model.	(Winkler et al. 2015)
Global	Partial dependence plot; individual conditional expectation plot; accumulated local effects plot	Shows the marginal effects of features on the prediction	Visualizes how the associations between the features and response are modeled.	(Hastie et al. 2001, Goldstein et al. 2014, Apley and Zhu 2019)
Global	H-statistic	Measures the interaction strength of paired features based on the stability of partial dependence	Identifies key pairwise interactions, either among all possible combinations or given a specific feature.	(Friedman and Popescu 2008)
Global	Greenwell's partial dependence-based variable importance	Measures the interaction strength of paired features based on the stability of partial dependence	Identifies key pairwise interactions, either among all possible combinations or given a specific feature.	(Greenwell et al. 2018)

as argmin $L(f, g, \pi_x) + \Omega(g)$. The term $L(f, g, \pi_x)$ calculates the difference in accuracy between the complex model f (e.g. random forests) and a simple model g (e.g. linear model) at the target prediction x and the surrounding neighborhood of proximity π in the n-dimensional space. The term $\Omega(g)$ is the complexity of the simple model represented as the number of parameters. The LIME algorithm minimizes $L + \Omega$ to replace the complex model by the simpler one, while attempting to avoid losing accuracy. A key assumption of LIME is that the necessary degree of model complexity depends on the data domain for which predictions should be made.

Hence, LIME helps us remove 'unnecessary' complexity from a global model to better understand how it arrives at local predictions. Although a complex algorithm may be necessary to accurately model species distributions at large spatial extents (e.g. the full species range), a simpler algorithm is often sufficiently accurate at smaller extents of the same grain size where conservation and management activities actually take place. In fact, many parameters that would apply to large spatial extents are not as important at small extents, where most of the parameters can often be assumed to be constants (but see Potter et al. 2013).

The local surrogate model can also be used for checking the ecological plausibility of model behavior and prediction, as we demonstrate in an example where we provide site-level assessment and interpretation for an SDM for the African elephant (Box 1, Fig. 2). Most complex algorithms were primarily designed to improve predictions, and design principles such as boosting, bagging or deep layers in neural networks usually complicate the interpretation of the fitted model. For example, suppose one fits a random forest model to a focal species with a range of different predictor variables and the model predicts the presence or high suitability for the species at a particular site. One may want to know why the model made such a prediction. For example, is it due to optimal climatic conditions, resource availability or other reasons?

Box 1. Explaining the distribution of the African elephant with xAI

We demonstrate here an application of some xAI tools from Table 1 for species distribution models (SDMs) with R (R Core Team), using the African elephant *Loxodonta africana* as an example. The R script to reproduce the analysis with detailed settings is available on Zenodo (<https://doi.org/10.5281/zenodo.4048271>). Note that our intention is purely demonstrational: we seek neither to advance the ecological knowledge of this species nor to adhere to all the best modeling practices (e.g. we did not consider spatial autocorrelation or model tuning).

We applied the random forests algorithm (Breiman 2001b) for modeling the distribution of *L. africana* using occurrence data downloaded from GBIF (Musila et al. 2019, naturgucker_de 2020, Navarro and Jackson 2020, Questagame 2020, Ueda 2020), 10 000 randomly sampled background points, and standard bioclimatic variables from WorldClim v2 (Fick and Hijmans 2017). For data acquisition and processing, we used the `sdmbench` package (Angelov 2018), for model training the `mlr` package (Bischl et al. 2016), and for model explanation the `iml` package (Molnar and Schratz 2020) and the `lime` package (Pedersen and Benesty 2019; but note that the `breakDown` package is an alternative, Biecek and Grudziaz 2020). The data was split into training and testing data, 70% and 30% respectively.

Conventionally, model assessment relies heavily on visual inspections of the mapped model predictions (in this case, species' habitat suitability; left panel in Fig. 2) and accuracy metrics. Besides these, xAI tools can be applied. In this example we apply permutation importance that can rank and identify important predictor variables, and partial dependence plots that can visualize how habitat suitability is associated with focal variables in the model (note that such techniques are also oftentimes used in conventional SDM assessment). We interpret that the model is accurate when evaluated on testing data (area under the ROC curve = 0.98) and that the most important variable is the precipitation of the wettest quarter. This variable is positively associated with habitat suitability to a certain level, but an excessive precipitation amount does not increase suitability. These model assessments are important for biogeographical understanding at the large spatial extent.

Moreover, using local surrogate approaches, we can evaluate the reliability of the model and the importance of variables at the local spatial extent, where actual management and/or conservation occurs. Such comprehensive assessments are rarely seen in the SDM literature. With local interpretable model-agnostic explanation (LIME), we show site-level model assessment at three randomly chosen sites (right panel). At site A, the model predicts high habitat suitability (1.00). With LIME, we can confirm that the prediction is supported by all top five environmental conditions at the site. At sites B and C, suitability is lower than site A. LIME suggests that precipitation during the driest season is unfavorable at site B, while temperature is unfavorable at site C (see variables of 'vote for absence' in the right panel).

While the global-level variable importance suggests precipitation of the wettest quarter as the most important variable, the variable is not listed as a key variable at the three sites. This may indicate that, while precipitation of the wettest quarter is the main variable that explains the variability of habitat suitability over the entire modeled region, other variables become more important (e.g. temperature seasonality) at the local scale.

Note that 1) we confirmed only top five variables for brevity and interpretation might change by considering all variables (but it is not recommended as it decreases interpretability), and 2) the local surrogate model did not perfectly explain the global model. We do not intend to solve these issues in this exercise, but they can be taken as potential caveats of LIME.

As demonstrated, individual LIME explanations for local sites can help us better explore spatial variations in variable importance, which in turn, can contribute to more reasonable conservation and management decisions with higher interpretability for the model at the local scale.

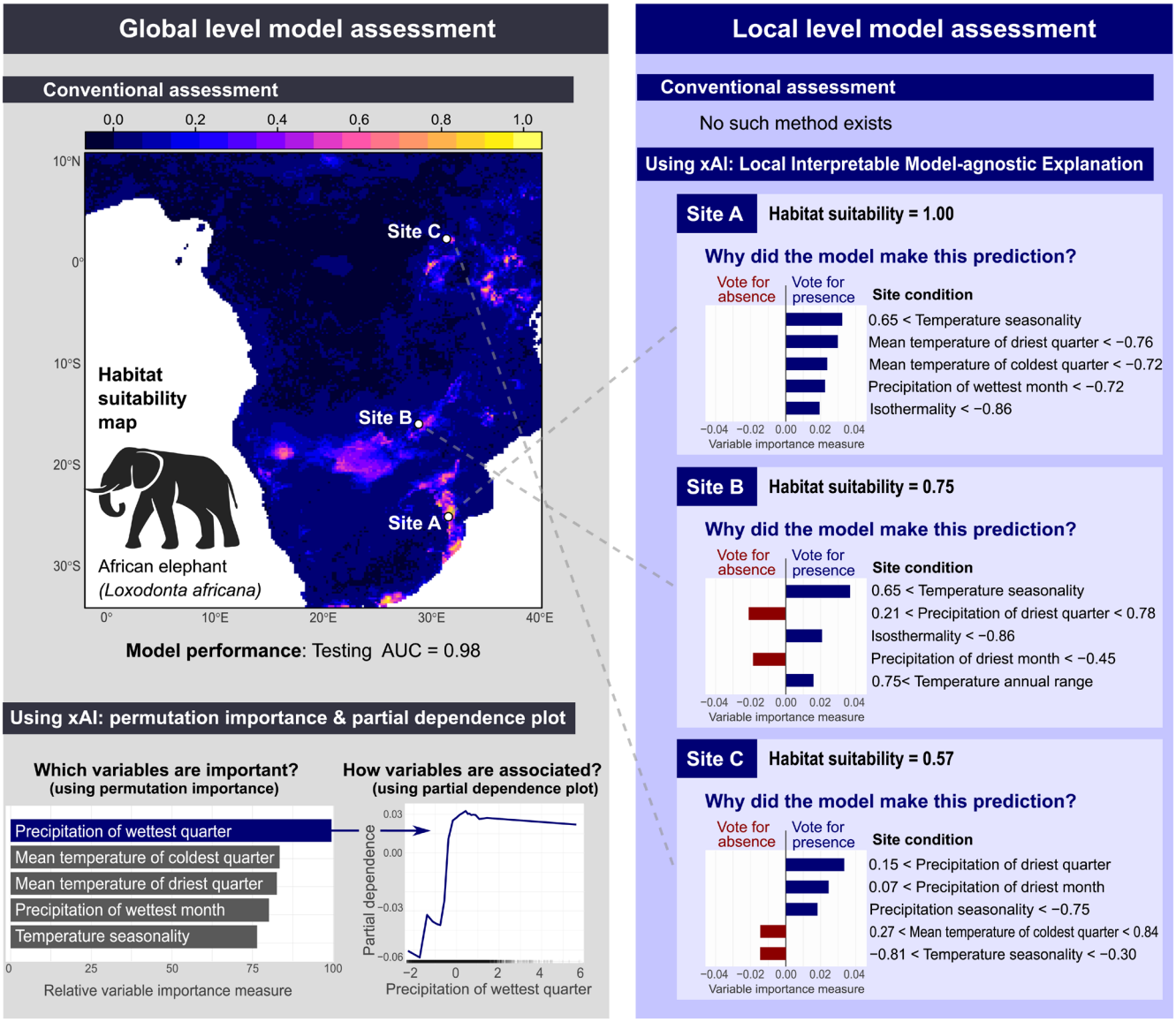

Figure 2. Interpreting the species distribution model of the African elephant *Loxodonta africana*, using explainable AI (xAI) tools: permutation importance and a partial dependence plot for the global level and local interpretable model-agnostic explanations at the local level (see Table 1 for the description of the techniques).

LIME can help to analyze how the importance of the predictor variables changes with scale and/or subregion (Ryo et al. 2018) and which variables are most relevant for a particular location or scenario.

More broadly, xAI methods can help researchers analyze and approximate the global and local behavior of the model and identify the reasons for why particular predictions are made (although important predictors may not be causally related to species' occurrences). It is widely appreciated that statistical models can use non-causal predictor variables to make predictions (i.e. the model predicts the right outcome for the wrong reason (Fourcade et al. 2018)). This is not necessarily a problem, because non-causal factors can act as proxies for unobserved and unobservable causal factors to improve predictions. However, the use of such non-causal model structures is problematic when predicting under conditions where the correlation structures of predictor variables change (Dormann et al. 2013). It is therefore important to determine the extent to which the fitted model reflects the true causal structure, and thus the mechanisms actually driving these relationships.

xAI cannot directly answer these questions, but it can help ecologists to examine the question of causality. For example, an xAI analysis may show that model predictions depend on predictor variables that are determined a priori as unlikely relevant for the focal species, or that the relevance of predictor variables changes in geographical or environmental space in a way that is ecologically counterintuitive. These results

may lead the researcher to reconsider the extent to which the fitted model reflects true mechanistic relationships, as well as the extent to which it can be used for extrapolation or to inform direct management interventions. Moreover, when using SDMs for extrapolation, for example to model species invasions, xAI tools can be applied to examine how variable importance changes from the fit to the extrapolation domain. In such a way, xAI can be combined with ecological and biogeographical knowledge to create a richer and more accurate interpretation of fitted machine learning models.

Whereas xAI would bring substantial benefits to the SDM research domain, we also acknowledge some major caveats. There is an ongoing debate on whether we should use post-hoc model-agnostic methods to explain complex models or use simpler models that can be more easily interpreted (Rudin 2019, Krishnan 2020, Molnar et al. 2020). As we also argued, 'explaining the modeled associations' is not the same as 'explaining the real causal associations'(Lipton 2018). Moreover, local surrogate approaches that explain the model behavior by another statistical model itself may bring additional error because of imperfect model approximation. For instance, LIME performance depends on parameter π, how the local domain for approximation is defined within the entire variable space. Certainly, interpretation cannot be made appropriately without domain expertise and a proper identification of the pitfalls associated with the method (Molnar et al. 2020).

In conclusion, we hope that this article will encourage applications of xAI tools in the SDM research domain to strengthen mutual understanding between modelers and practitioners. Post-hoc model-agnostic tools we list in Table 1 can be used for virtually any SDM algorithm, including popular algorithms such as Maxent. Due to its high compatibility with existing modeling approaches, xAI can be used to compare multiple models. Expert knowledge from both groups can be used to assess how local predictions are made based on the output of xAI, and this should inform model selection and conservation or management action. To conclude, we think that demystifying the decisions that complex models make is a necessary step towards producing models that can explain real-world ecological phenomena (Araújo et al. 2019, Mammola et al. 2019).

Data availability statement

Data available from the Zenodo Digital Repository: <https://doi.org/10.5281/zenodo.4048271> (Angelov 2020).

Acknowledgements – We thank Matthias C. Rillig and two anonymous reviewers for helping to improve the clarity of the manuscript. Open access funding enabled and organized by Projekt DEAL.

Funding – MR acknowledges funding by the Grant-in-Aid for JSPS Overseas Research Fellowships. BMB acknowledges support from the European Research Council (BIODESERT project, ERC Grant agreement no. 647038). SM acknowledges support from the European Commission through Horizon 2020 Marie Skłodowska-Curie Actions (MSCA) individual fellowships (Grant no. 882221). JMK acknowledges support from the Japan Society for the Promotion of Science Postdoctoral Fellowships for Foreign Researchers program and the Okinawa Institute of Science and Technology Graduate University. We also acknowledge support by the OpenAccess Publication Fund of Freie Universität Berlin.

Author contributions – MR and BA conceived the idea, designed the methodology and wrote the R script and analysed the data. SM, JMK, BMB and FH provided substantial suggestions on species distribution modeling and other theoretical aspects. MR wrote the first draft of the paper and all authors contributed to the writing with suggestions and comments.

References

Adadi, A. and Berrada, M. 2018. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). – IEEE Access 6: 52138–52160.

Angelov, B. 2018. sdmbench: R package for benchmarking species distribution models. – J. Open Source Softw. 3: 847.

Angelov, B. 2019. Review of species distribution modeling open-source software. – Researchgate doi: 10.13140/RG.2.2.33804.67203.

Angelov, B. 2020. boyanangelov/interpretable_sdm: reproducibility fix (ver. v1.3). – Zenodo, <http://doi.org/10.5281/zenodo.4048271>.

Apley, D. W. and Zhu, J. 2019. Visualizing the effects of predictor variables in black box supervised learning models. – arXiv: 1612.08468 [stat.ME].

Araújo, M. B. et al. 2019. Standards for distribution models in biodiversity assessments. – Sci. Adv. 5: eaat4858.

Austin, M. 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. – Ecol. Model. 200: 1–19.

Biecek, P. and Grudziaz, A. 2020. breakDown: model agnostic explainers for individual predictions. – <https://cran.r-project.org/web/packages/breakDown/breakDown.pdf>.

Bischl, B. et al. 2016. mlr: machine learning in R. – J. Mach. Learn. Res. 17: 1–5.

Boehmke, B. and Greenwell, B. 2020. Hands-on machine learning with R. – <https://bradleyboehmke.github.io/HOML/>.

Booth, T. H. 2018. Why understanding the pioneering and continuing contributions of BIOCLIM to species distribution modelling is important. – Austral Ecol. 43: 852–860.

Booth, T. H. et al. 2014. bioclim: the first species distribution modelling package, its early applications and relevance to most current MaxEnt studies. – Divers. Distrib. 20: 1–9.

Box, E. O. 1981. Predicting physiognomic vegetation types with climate variables. – Vegetatio 45: 127–139.

Breiman, L. 2001a. Statistical modeling: the two cultures. – Stat. Sci. 16: 199–215.

Breiman, L. 2001b. Random Forests. – Mach. Lang. 45: 5–32.

Carvalho, D. V. et al. 2019. Machine learning interpretability: a survey on methods and metrics. – Electronics 8: 832.

Cutler, D. R. et al. 2007. Random forests for classification in ecology. – Ecology 88: 2783–2792.

de Souza Muñoz, M. E. et al. 2011. openModeller: a generic approach to species' potential distribution modelling. – GeoInformatica 15: 111–135.

Dormann, C. F. et al. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. – Ecography 36: 27–46.

Elith, J. and Graham, C. H. 2009. Do they? How do they? Why do they differ? On finding reasons for differing performances of species distribution models. – Ecography 32: 66–77.
Elith, J. and Leathwick, J. R. 2009. Species distribution models: ecological explanation and prediction across space and time. – Annu. Rev. Ecol. Evol. Syst. 40: 677–697.
Elith, J. et al. 2008. A working guide to boosted regression trees. – J. Anim. Ecol. 77: 802–813.
Elith, J. et al. 2011. A statistical explanation of MaxEnt for ecologists. – Divers. Distrib. 17: 43–57.
Fick, S. E. and Hijmans, R. J. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. – Int. J. Climatol. 37: 4302–4315.
Fourcade, Y. et al. 2018. Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. – Global Ecol. Biogeogr. 27: 245–256.
Friedman, J. H. and Popescu, B. E. 2008. Predictive learning via rule ensembles. – Ann. Appl. Stat. 2: 916–954.
Goldstein, A. et al. 2014. Peeking inside the black box: visualizing statistical learning with plots of individual conditional expectation. – arXiv:1309.6392 [stat.AP].
Gosiewska, A. and Biecek, P. 2020. Do not trust additive explanations. – arXiv:1903.11420 [cs.LG].
Greenwell, B. M. et al. 2018. A simple and effective model-based variable importance measure. – arXiv:1805.04755 [stat.ML].
Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – Ecol. Lett. 8: 993–1009.
Halvorsen, R. et al. 2015. Opportunities for improved distribution modelling practice via a strict maximum likelihood interpretation of MaxEnt. – Ecography 38: 172–183.
Hastie, T. et al. 2001. The elements of statistical learning: data mining, inference and prediction. – Springer Science and Business Media.
Kass, J. M. et al. 2018. Wallace: a flexible platform for reproducible modeling of species niches and distributions built for community expansion. – Methods Ecol. Evol. 9: 1151–1156.
Krishnan, M. 2020. Against interpretability: a critical examination of the interpretability problem in machine learning. – Phil. Technol. 33: 487–502.
Lipton, Z. C. 2018. In machine learning, the concept of interpretability is both important and slippery. – Queue 16: 28.
Lucas, T. C. D. 2020. A translucent box: interpretable machine learning in ecology. – Ecol. Monogr. 90: e01422.
Lundberg, S. and Lee, S.-I. 2017. A unified approach to interpreting model predictions. – arXiv:1705.07874 [cs.AI].
Mammola, S. et al. 2019. Associations between habitat quality, body size and reproductive fitness in the alpine endemic spider *Vesubia jugorum*. – Global Ecol. Biogeogr. 28: 1325–1335.
Merow, C. et al. 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. – Ecography 36: 1058–1069.
Merow, C. et al. 2014. What do we gain from simplicity versus complexity in species distribution models? – Ecography 37: 1267–1281.
Meske, C. and Bunde, E. 2020. Using explainable artificial intelligence to increase trust in computer vision. – arXiv:cs.HC/2002.01543.
Molnar, C. 2019. Interpretable machine learning. – <https://christophm.github.io/interpretable-ml-book/>.
Molnar, C. and Schratz, P. 2020. iml: interpretable machine learning. – J. Open Source Softw. 3: 786.
Molnar, C. et al. 2020. Pitfalls to avoid when interpreting machine learning models. – arXiv:2007.04131 [stat.ML].
Murdoch, W. J. et al. 2019. Definitions, methods and applications in interpretable machine learning. – Proc. Natl Acad. Sci. USA 116: 22071–22080.
Musila, S. et al. 2019. Occurrence records of mammal species in Tana River Basin, Kenya v1.1. – Natl Mus. Kenya.
naturgucker_de 2020. naturgucker. – <www.gbif.org/dataset/6ac3f774-d9fb-4796-b3e9-92bf6c81c084>, accessed 30 May 2020.
Navarro, R. and Jackson, C. 2020. Kenya virtual museum records v1.11. – <www.gbif.org/dataset/999d8fed-d06d-48d8-87ee-cc36e6c8456c>, accessed 30 May 2020.
Olden, J. D. and Jackson, D. A. 2002. Illuminating the 'black box': a randomization approach for understanding variable contributions in artificial neural networks. – Ecol. Model. 154: 135–150.
Olden, J. D. et al. 2008. Machine learning methods without tears: a primer for ecologists. – Q. Rev. Biol. 83: 171–193.
Pecl, G. T. et al. 2017. Biodiversity redistribution under climate change: impacts on ecosystems and human well-being. – Science 355: eaai9214.
Pedersen, T. L. and Benesty, M. 2019. lime: local interpretable model-agnostic explanations. – <https://cran.r-project.org/web/packages/lime/lime.pdf>.
Phillips, S. J. 2017. A brief tutorial on Maxent. – <http://biodiversityinformatics.amnh.org/open_source/maxent/>, accessed 21 June 2020.
Phillips, S. J. et al. 2004. A maximum entropy approach to species distribution modeling. – In: Proc. Twenty-First Int. Conf. Mach. Learn., pp. 655–662.
Phillips, S. J. et al. 2006. Maximum entropy modeling of species geographic distributions. – Ecol. Model. 190: 231–259.
Potter, K. A. et al. 2013. Microclimatic challenges in global change biology. – Global Change Biol. 19: 2932–2939.
Qiao, H. et al. 2015. No silver bullets in correlative ecological niche modelling: insights from testing among many potential algorithms for niche estimation. – Methods Ecol. Evol. 6: 1126–1136.
Questagame. 2020. Questagame weekly feed. – <www.gbif.org/dataset/e3ce628e-9683-4af7-b7a9-47eef785d3bb/metrics>, accessed 30 May 2020.
Ribeiro, M. T. et al. 2016. 'Why should I trust you?': explaining the predictions of any classifier. – In: Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., pp. 1135–1144.
Ribeiro, M. T. et al. 2018. Anchors: high-precision model-agnostic explanations. – AAAI, pp. 1527–1535.
Rudin, C. 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. – Nat. Mach. Intell. 1: 206–215.
Ryo, M. and Rillig, M. C. 2017. Statistically reinforced machine learning for nonlinear patterns and variable interactions. – Ecosphere 8: e01976.
Ryo, M. et al. 2018. Nonlinear higher order abiotic interactions explain riverine biodiversity. – J. Biogeogr. 45: 628–639.
Scachetti-Pereira, R. 2002. DesktopGarp: a software package for biodiversity and ecologic research. – The Univ. of Kansas Biodiversity Research Center, USA, <www.nhm.ku.edu/desktopgarp/>.
Ueda, K. 2020. iNaturalist research-grade observations. – <www.gbif.org/dataset/50c9509d-22c7-4a22-a47d-8c48425ef4a7>, accessed 30 May 2020.
Winkler, A. M. et al. 2015. Multi-level block permutation. – NeuroImage 123: 253–268.